Self-Interested Agents and Utility Theory

Game Theory Course:
Jackson, Leyton-Brown & Shoham

Self-interested agents

- What does it mean to say that an agent is self-interested?
  - not that they want to harm others or only care about themselves
  - only that the agent has its own description of states of the world that it likes, and acts based on this description

Self-interested agents

- What does it mean to say that an agent is self-interested?
  - not that they want to harm others or only care about themselves
  - only that the agent has its own description of states of the world that it likes, and acts based on this description

- Each such agent has a utility function
  - quantifies degree of preference across alternatives
  - explains the impact of uncertainty
  - Decision-theoretic rationality: act to maximize expected utility.

Why Utility?

- It might seem obvious that preferences can be described by utility functions. But:
  - Why is a single-dimensional function enough?
  - Why should an agent's response to uncertainty be captured purely by an expected value?
- Thus, our claim is substantive.

- There's a famous theorem (von Neumann & Morgenstern, 1944) that derives the existence of a utility function from a more basic preference ordering and axioms on such orderings.